

## RAW SEQUENCE LISTING ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 10/729,571  
Source: IFW  
Date Processed by STIC: 12/22/03

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.

PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

FOR CRF SUBMISSION AND PATENTIN SOFTWARE QUESTIONS, PLEASE CONTACT MARK SPENCER, TELEPHONE: 703-308-4212; FAX: 703-308-4221

Effective 12/13/03: TELEPHONE: 571-272-2510; FAX: 571-273-0221

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER VERSION 4.1 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:

<http://www.uspto.gov/web/offices/pac/checker/chkr41note.htm>

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.

Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom.

Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:

1. EFS-Bio (<<http://www.uspto.gov/ebc/efs/downloads/documents.htm>> , EFS Submission User Manual - ePAVE)
2. U.S. Postal Service: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450
3. Hand Carry directly to (EFFECTIVE 12/01/03):  
U.S. Patent and Trademark Office, Box Sequence, Customer Window, Lobby, Room 1B03, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202
4. Federal Express, United Parcel Service, or other delivery service to: U.S. Patent and Trademark Office, Box Sequence, Room 4B03-Mailroom, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202

Revised 10/08/03



/Fwo

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/729,571

DATE: 12/22/2003

TIME: 10:11:23

Input Set : A:\SEQUENCE LISTING.txt

Output Set: N:\CRF4\12222003\J729571.raw

3 <110> APPLICANT: Anderson et al.  
 5 <120> TITLE OF INVENTION: CRYSTAL STRUCTURE OF GLUTAMATE RACEMASE (MURL)  
 7 <130> FILE REFERENCE: ASZD-P01-007  
 C--> 9 <140> CURRENT APPLICATION NUMBER: US/10/729,571  
 C--> 10 <141> CURRENT FILING DATE: 2003-12-05  
 12 <150> PRIOR APPLICATION NUMBER: 60/435,272  
 13 <151> PRIOR FILING DATE: 2002-12-20  
 15 <150> PRIOR APPLICATION NUMBER: 60/435,167  
 16 <151> PRIOR FILING DATE: 2002-12-20  
 18 <150> PRIOR APPLICATION NUMBER: 60/435,087  
 19 <151> PRIOR FILING DATE: 2002-12-20  
 21 <150> PRIOR APPLICATION NUMBER: 60/435,527  
 22 <151> PRIOR FILING DATE: 2002-12-20  
 24 <160> NUMBER OF SEQ ID NOS: 76  
 26 <170> SOFTWARE: PatentIn version 3.2

*error throughout*  
 Does Not Comply  
 Corrected Diskette Needed

## ERRORED SEQUENCES

28 <210> SEQ ID NO: 1  
 29 <211> LENGTH: 768  
 30 <212> TYPE: DNA  
 31 <213> ORGANISM: H. pylori  
 33 <221> NAME/KEY: CDS  
 34 <222> LOCATION: (1)..(768)  
 OK> 36 <400> SEQUENCE: 1

37	atg	aaa	ata	ggc	gtt	ttt	gat	agc	ggt	gtg	ggg	ggg	ttt	agc	gtt	tta	48
38	Met	Lys	Ile	Gly	Val	Phe	Asp	Ser	Gly	Val	Gly	Gly	Phe	Ser	Val	Leu	
39	1			5					10				15				
41	aaa	agc	ctt	tta	aaa	gcg	cga	ttg	ttt	gat	gaa	atc	atc	tac	tat	ggc	96
42	Lys	Ser	Leu	Leu	Lys	Ala	Arg	Leu	Phe	Asp	Glu	Ile	Ile	Tyr	Tyr	Gly	
43			20					25					30				
45	gat	agc	gct	aga	gtg	cct	tat	ggc	act	aaa	gac	ccc	acc	acg	atc	aag	144
46	Asp	Ser	Ala	Arg	Val	Pro	Tyr	Gly	Thr	Lys	Asp	Pro	Thr	Thr	Ile	Lys	
47			35					40				45					
49	caa	ttt	ggc	tta	gag	gct	ttg	gat	ttt	ttc	aaa	cgc	cat	gag	att	gaa	192
50	Gln	Phe	Gly	Leu	Glu	Ala	Leu	Asp	Phe	Phe	Lys	Pro	His	Glu	Ile	Glu	
51		50					55					60					
53	tta	ttg	att	gtg	gca	tgc	aac	acc	gcg	agc	gct	ctg	gct	tta	gaa	gag	240
54	Leu	Leu	Ile	Val	Ala	Cys	Asn	Thr	Ala	Ser	Ala	Leu	Ala	Leu	Glu	Glu	
55	65			70				75				80					
57	atg	caa	aag	tat	tct	aaa	atc	cct	att	gtg	ggc	gtg	att	gag	cca	agc	288
58	Met	Gln	Lys	Tyr	Ser	Lys	Ile	Pro	Ile	Val	Gly	Val	Ile	Glu	Pro	Ser	

*(global error)*  
 insert this mandatory numeric identifier  
 whenever <221>, <222>, or <223> is  
 shown

## RAW SEQUENCE LISTING

DATE: 12/22/2003

PATENT APPLICATION: US/10/729,571

TIME: 10:11:23

Input Set : A:\SEQUENCE LISTING.txt

Output Set: N:\CRF4\12222003\J729571.raw

```

59          85          90          95
61 att tta gcg atc aag cgg caa gtg gaa gat aaa aac gcc cct att tta      336
62 Ile Leu Ala Ile Lys Arg Gln Val Glu Asp Lys Asn Ala Pro Ile Leu
63          100          105          110
65 gtg cta ggg aca aaa gcg acg att caa tcc aac gcc tat gac aac gcc      384
66 Val Leu Gly Thr Lys Ala Thr Ile Gln Ser Asn Ala Tyr Asp Asn Ala
67          115          120          125
69 ctg aaa caa caa ggc tat ttg aac att tcg cat tta gct act tct ctt      432
70 Leu Lys Gln Gln Gly Tyr Leu Asn Ile Ser His Leu Ala Thr Ser Leu
71          130          135          140
73 ttt gtg cct ttg att gaa gaa agt att tta gag ggc gaa ttg tta gaa      480
74 Phe Val Pro Leu Ile Glu Glu Ser Ile Leu Glu Gly Glu Leu Leu Glu
75 145          150          155          160
77 act tgc atg cat tat tat ttc act ccc tta gag att tta ccc gaa gtg      528
78 Thr Cys Met His Tyr Phe Thr Pro Leu Glu Ile Leu Pro Glu Val
79          165          170          175
81 atc att tta ggt tgc acg cat ttt ccc tta atc gct caa aaa att gag      576
82 Ile Ile Leu Gly Cys Thr His Phe Pro Leu Ile Ala Gln Lys Ile Glu
83          180          185          190
85 ggc tat ttc atg ggg cat ttt gcc ctt cca acg ccc ccc cta ctc atc      624
86 Gly Tyr Phe Met Gly His Phe Ala Leu Pro Thr Pro Pro Leu Leu Ile
87          195          200          205
89 cat tcg ggc gat gct att gta gaa tat ttg caa caa aaa tac gcc ctt      672
90 His Ser Gly Asp Ala Ile Val Glu Tyr Leu Gln Gln Lys Tyr Ala Leu
91          210          215          220
93 aaa aac aat gca tgc aca ttc cct aaa gtg gaa ttt cat gcg agc ggc      720
94 Lys Asn Asn Ala Cys Thr Phe Pro Lys Val Glu Phe His Ala Ser Gly
95 225          230          235          240
97 gat gtg atc tgg cta gaa aga caa gct aaa gaa tgg ctc aaa ttg taa      768
98 Asp Val Ile Trp Leu Glu Arg Gln Ala Lys Glu Trp Leu Lys Leu
99          245          250          255

```

171 &lt;210&gt; SEQ ID NO: 3

172 &lt;211&gt; LENGTH: 768

173 &lt;212&gt; TYPE: DNA

174 <213> ORGANISM: *H. pylori*

176 &lt;221&gt; NAME/KEY: CDS

177 &lt;222&gt; LOCATION: (1)..(768)

✶-&gt; 179 &lt;400&gt; SEQUENCE: 3

```

180 atg aaa ata ggc gtt ttt gat agc ggt gtg ggg ggg ttt agc gtt tta      48
181 Met Lys Ile Gly Val Phe Asp Ser Gly Val Gly Gly Phe Ser Val Leu
182 1          5          10          15
184 aaa agc ctt tta aaa gcg caa ttg ttt gat gaa atc atc tat tat ggc      96
185 Lys Ser Leu Leu Lys Ala Gln Leu Phe Asp Glu Ile Ile Tyr Tyr Gly
186          20          25          30
188 gat agc gct aga gtg cct tat ggc act aaa gac ccc act acg atc aag      144
189 Asp Ser Ala Arg Val Pro Tyr Gly Thr Lys Asp Pro Thr Thr Ile Lys
190          35          40          45
192 caa ttt ggc tta gag gct ttg gat ttt ttc aaa cca cac cag att gaa      192
193 Gln Phe Gly Leu Glu Ala Leu Asp Phe Phe Lys Pro His Gln Ile Glu

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/729,571

DATE: 12/22/2003

TIME: 10:11:23

Input Set : A:\SEQUENCE LISTING.txt

Output Set: N:\CRF4\12222003\J729571.raw

```

194      50      55      60
196 tta ttg att gtg gca tgc aac acc gca agc gct ctg gct tta gaa gag      240
197 Leu Leu Ile Val Ala Cys Asn Thr Ala Ser Ala Leu Ala Leu Glu Glu
198 65      70      75      80
200 atg caa aag cat tcc aaa atc cct att gtg ggc gtg att gag cca agc      288
201 Met Gln Lys His Ser Lys Ile Pro Ile Val Gly Val Ile Glu Pro Ser
202      85      90      95
204 att tta gcg atc aag caa caa gtg aaa gat aaa aac gcc cct att tta      336
205 Ile Leu Ala Ile Lys Gln Gln Val Lys Asp Lys Asn Ala Pro Ile Leu
206      100      105      110
208 gtg cta ggg aca aaa gcg acg att caa tcc aac gct tat gac aac gcc      384
209 Val Leu Gly Thr Lys Ala Thr Ile Gln Ser Asn Ala Tyr Asp Asn Ala
210      115      120      125
212 ctg aaa caa caa ggc tat ttg aat gtt tcg cat tta gcc act tct ctt      432
213 Leu Lys Gln Gln Gly Tyr Leu Asn Val Ser His Leu Ala Thr Ser Leu
214      130      135      140
216 ttt gtg cct ttg att gaa gaa agt att tta gag ggc gaa ttg tta gag      480
217 Phe Val Pro Leu Ile Glu Glu Ser Ile Leu Glu Gly Glu Leu Leu Glu
218 145      150      155      160
220 act tgc atg cgt tat tat ttc act ccc tta aag att tta cct gaa gtg      528
221 Thr Cys Met Arg Tyr Tyr Phe Thr Pro Leu Lys Ile Leu Pro Glu Val
222      165      170      175
224 att att tta ggt tgc acg cat ttt ccc ttg att gct caa aaa att gag      576
225 Ile Ile Leu Gly Cys Thr His Phe Pro Leu Ile Ala Gln Lys Ile Glu
226      180      185      190
228 ggc tat ttc atg gag cat ttt gcc ctt cca acg ccc ccc cta ctc atc      624
229 Gly Tyr Phe Met Glu His Phe Ala Leu Pro Thr Pro Pro Leu Leu Ile
230      195      200      205
232 cat tcg ggc gat gct att gta gaa tat ttg cag caa aaa tac gcc ctt      672
233 His Ser Gly Asp Ala Ile Val Glu Tyr Leu Gln Gln Lys Tyr Ala Leu
234      210      215      220
236 aaa aac aat gca cac gca ttc cct aaa gtg gaa ttt cat gcg agc ggc      720
237 Lys Asn Asn Ala His Ala Phe Pro Lys Val Glu Phe His Ala Ser Gly
238 225      230      235      240
240 gat gtg atc tgg cta gaa aga caa gct aaa gaa tgg ctc aaa ttg taa      768
241 Asp Val Ile Trp Leu Glu Arg Gln Ala Lys Glu Trp Leu Lys Leu
242      245      250      255
314 <210> SEQ ID NO: 5
315 <211> LENGTH: 768
316 <212> TYPE: DNA
317 <213> ORGANISM: H. pylori
319 <221> NAME/KEY: CDS
320 <222> LOCATION: (1)..(768)
322 <400> SEQUENCE: 5
323 atg aaa ata ggc gtt ttt gat agc ggt gtg gga ggg ttt agc gtt tta      48
324 Met Lys Ile Gly Val Phe Asp Ser Gly Val Gly Gly Phe Ser Val Leu
325 1      5      10      15
327 aaa agc ctt tta aaa gcg caa tta ttt gat gaa atc atc tat tat ggc      96
328 Lys Ser Leu Leu Lys Ala Gln Leu Phe Asp Glu Ile Ile Tyr Tyr Gly

```

④&lt;-&gt;

&gt;2207

## RAW SEQUENCE LISTING

DATE: 12/22/2003

PATENT APPLICATION: US/10/729,571

TIME: 10:11:23

Input Set : A:\SEQUENCE LISTING.txt

Output Set : N:\CRF4\12222003\J729571.raw

```

329          20          25          30
331 gat agc gct aga gtg cct tat ggc act aaa gac ccc act acg atc aag      144
332 Asp Ser Ala Arg Val Pro Tyr Gly Thr Lys Asp Pro Thr Thr Ile Lys
333          35          40          45
335 caa ttt ggc tta gag gct ttg gat ttt ttc aaa ccg cac cag att gaa      192
336 Gln Phe Gly Leu Glu Ala Leu Asp Phe Phe Lys Pro His Gln Ile Glu
337          50          55          60
339 tta ttg att gtg gca tgc aac aca gcg agc gct cta gct tta gaa gag      240
340 Leu Leu Ile Val Ala Cys Asn Thr Ala Ser Ala Leu Ala Leu Glu Glu
341 65          70          75          80
343 atg caa aag cat tcc aaa atc cct att gtg ggc gtg att gag cca agc      288
344 Met Gln Lys His Ser Lys Ile Pro Ile Val Gly Val Ile Glu Pro Ser
345          85          90          95
347 att tta gcg atc aag cga caa gta aaa gat aaa aac gcc cct att tta      336
348 Ile Leu Ala Ile Lys Arg Gln Val Lys Asp Lys Asn Ala Pro Ile Leu
349          100          105          110
351 gtg cta ggg aca aaa gcg acg atc caa tcc aac gct tat gac aat gcc      384
352 Val Leu Gly Thr Lys Ala Thr Ile Gln Ser Asn Ala Tyr Asp Asn Ala
353          115          120          125
355 ctg aaa caa caa ggc tat ttg aat gtt tcg cat tta gcc act tct ctt      432
356 Leu Lys Gln Gln Gly Tyr Leu Asn Val Ser His Leu Ala Thr Ser Leu
357          130          135          140
359 ttt gtg cct ttg att gaa gaa agt att tta gag ggc gaa ttg tta gaa      480
360 Phe Val Pro Leu Ile Glu Glu Ser Ile Leu Glu Gly Glu Leu Leu Glu
361 145          150          155          160
363 act tgc atg cgt tat tat ttc act ccc tta aag att tta ccc gaa gtg      528
364 Thr Cys Met Arg Tyr Tyr Phe Thr Pro Leu Lys Ile Leu Pro Glu Val
365          165          170          175
367 att att tta ggt tgc acg cat ttt ccc tta atc gct caa aaa att gag      576
368 Ile Ile Leu Gly Cys Thr His Phe Pro Leu Ile Ala Gln Lys Ile Glu
369          180          185          190
371 ggc tat ttt atg gag cat ttt gcc ctt tca aca ccc ccc cta ctc atc      624
372 Gly Tyr Phe Met Glu His Phe Ala Leu Ser Thr Pro Pro Leu Leu Ile
373          195          200          205
375 cat tcg ggc gat gct att gta gga tat ttg cag caa aaa tac gcc ctt      672
376 His Ser Gly Asp Ala Ile Val Gly Tyr Leu Gln Gln Lys Tyr Ala Leu
377          210          215          220
379 aaa aaa aat gca cac gca ttc cct aaa gtg gaa ttt cat gcg agc ggc      720
380 Lys Lys Asn Ala His Ala Phe Pro Lys Val Glu Phe His Ala Ser Gly
381 225          230          235          240
383 gat gtg atc tgg cta gaa aaa caa gct aaa gaa tgg ctc aaa ttg taa      768
384 Asp Val Ile Trp Leu Glu Lys Gln Ala Lys Glu Trp Leu Lys Leu
385          245          250          255

```

457 &lt;210&gt; SEQ ID NO: 7

458 &lt;211&gt; LENGTH: 749

459 &lt;212&gt; TYPE: DNA

460 <213> ORGANISM: *H. pylori*

462 &lt;221&gt; NAME/KEY: CDS

463 &lt;222&gt; LOCATION: (1)..(747)

-&gt;&lt;220&gt;

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/729,571

DATE: 12/22/2003

TIME: 10:11:23

Input Set : A:\SEQUENCE LISTING.txt

Output Set: N:\CRF4\12222003\J729571.raw

OK-> 465 <400> SEQUENCE: 7

466	atg	aaa	ata	ggc	gtt	ttt	gat	agc	ggt	gtg	gga	ggg	ttt	agc	gtt	tta	48
467	Met	Lys	Ile	Gly	Val	Phe	Asp	Ser	Gly	Val	Gly	Gly	Phe	Ser	Val	Leu	
468	1			5					10				15				
470	aaa	agc	ctt	tta	aaa	gcg	caa	ttg	ttt	gat	gaa	atc	atc	tat	tat	ggc	96
471	Lys	Ser	Leu	Leu	Lys	Ala	Gln	Leu	Phe	Asp	Glu	Ile	Ile	Tyr	Tyr	Gly	
472			20					25					30				
474	gat	agc	gct	aga	gtg	cct	tat	ggc	act	aaa	gac	ccc	acc	acg	atc	aag	144
475	Asp	Ser	Ala	Arg	Val	Pro	Tyr	Gly	Thr	Lys	Asp	Pro	Thr	Thr	Ile	Lys	
476			35					40				45					
478	caa	ttt	ggc	tta	gag	gct	ttg	gat	ttt	ttc	aaa	ccg	cac	cag	att	aaa	192
479	Gln	Phe	Gly	Leu	Glu	Ala	Leu	Asp	Phe	Phe	Lys	Pro	His	Gln	Ile	Lys	
480		50					55				60						
482	tta	ttg	att	gtg	gca	tgc	aac	aca	gcg	agc	gct	cta	gct	tta	gaa	gag	240
483	Leu	Leu	Ile	Val	Ala	Cys	Asn	Thr	Ala	Ser	Ala	Leu	Ala	Leu	Glu	Glu	
484	65			70					75			80					
486	atg	caa	aag	cat	tcc	aaa	atc	cct	att	gtg	ggc	gtg	att	gag	cca	agc	288
487	Met	Gln	Lys	His	Ser	Lys	Ile	Pro	Ile	Val	Gly	Val	Ile	Glu	Pro	Ser	
488			85						90				95				
490	att	tta	gcg	atc	aag	caa	caa	gta	aaa	gat	aaa	aac	gcc	cct	att	tta	336
491	Ile	Leu	Ala	Ile	Lys	Gln	Gln	Val	Lys	Asp	Lys	Asn	Ala	Pro	Ile	Leu	
492			100					105				110					
494	gtg	cta	ggg	aca	aaa	gcg	acg	atc	caa	tcc	aac	gct	tat	gac	aac	gcc	384
495	Val	Leu	Gly	Thr	Lys	Ala	Thr	Ile	Gln	Ser	Asn	Ala	Tyr	Asp	Asn	Ala	
496			115				120				125						
498	ctg	aaa	caa	caa	ggc	tat	ttg	aat	gtt	tcg	cat	tta	gcc	act	tct	ctt	432
499	Leu	Lys	Gln	Gln	Gly	Tyr	Leu	Asn	Val	Ser	His	Leu	Ala	Thr	Ser	Leu	
500		130				135					140						
502	ttt	gtg	cct	ttg	att	gaa	gaa	agt	att	tta	ggg	ggc	gaa	ttg	tta	gaa	480
503	Phe	Val	Pro	Leu	Ile	Glu	Glu	Ser	Ile	Leu	Gly	Gly	Glu	Leu	Leu	Glu	
504	145			150					155			160					
506	act	tgc	atg	cgt	tat	ttc	act	ccc	tta	aag	att	tta	cct	gaa	gtg		528
507	Thr	Cys	Met	Arg	Tyr	Phe	Thr	Pro	Leu	Lys	Ile	Leu	Pro	Glu	Val		
508			165					170				175					
510	att	att	tta	ggt	tgc	acg	cat	ttt	ccc	ttg	atc	gct	caa	aaa	att	gag	576
511	Ile	Ile	Leu	Gly	Cys	Thr	His	Phe	Pro	Leu	Ile	Ala	Gln	Lys	Ile	Glu	
512			180				185				190						
514	ggc	tat	ttt	atg	gag	cat	ttt	gcc	ctt	tca	acg	ccc	ccc	cta	ctc	atc	624
515	Gly	Tyr	Phe	Met	Glu	His	Phe	Ala	Leu	Ser	Thr	Pro	Pro	Leu	Leu	Ile	
516			195				200				205						
518	cat	tcg	ggc	gat	gct	att	gtg	gaa	tat	ttg	cag	caa	aaa	tac	gcc	ctt	672
519	His	Ser	Gly	Asp	Ala	Ile	Val	Glu	Tyr	Leu	Gln	Gln	Lys	Tyr	Ala	Leu	
520		210				215				220							
522	aag	aaa	aat	gca	cac	gca	ttc	cct	aaa	gtg	gaa	ttt	cat	gcg	agc	ggc	720
523	Lys	Lys	Asn	Ala	His	Ala	Phe	Pro	Lys	Val	Glu	Phe	His	Ala	Ser	Gly	
524	225			230				235				240					
526	gat	gtg	atc	tgg	cta	gaa	aaa	cag	gct	aa							749
527	Asp	Val	Ile	Trp	Leu	Glu	Lys	Gln	Ala								
528				245													

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/729,571

DATE: 12/22/2003

TIME: 10:11:23

Input Set : A:\SEQUENCE LISTING.txt

Output Set: N:\CRF4\12222003\J729571.raw

600 <210> SEQ ID NO: 9  
 601 <211> LENGTH: 768  
 602 <212> TYPE: DNA  
 603 <213> ORGANISM: H. pylori  
 605 <221> NAME/KEY: CDS  
 606 <222> LOCATION: (1)..(768)

OK-> 608 <400> SEQUENCE: 9

609 atg aaa ata ggc gtt ttt gat agc ggt gtg gga ggg ttt agc gtt tta 48  
 610 Met Lys Ile Gly Val Phe Asp Ser Gly Val Gly Gly Phe Ser Val Leu  
 611 1 5 10 15  
 613 aaa agc ctt tta aaa gcg caa cta ttt gat gaa atc atc tat tat ggc 96  
 614 Lys Ser Leu Leu Lys Ala Gln Leu Phe Asp Glu Ile Ile Tyr Tyr Gly  
 615 20 25 30  
 617 gat agt gct aga gtg cct tat ggc act aaa gac ccc acc acg atc aag 144  
 618 Asp Ser Ala Arg Val Pro Tyr Gly Thr Lys Asp Pro Thr Thr Ile Lys  
 619 35 40 45  
 621 caa ttt ggc tta gag gct ttg gat ttt ttc aaa ccg cac cag att gga 192  
 622 Gln Phe Gly Leu Glu Ala Leu Asp Phe Phe Lys Pro His Gln Ile Gly  
 623 50 55 60  
 625 tta ttg att gtg gca tgc aac aca gcg agc gct cta gct tta gaa gag 240  
 626 Leu Leu Ile Val Ala Cys Asn Thr Ala Ser Ala Leu Ala Leu Glu Glu  
 627 65 70 75 80  
 629 atg caa aag cat tcc aaa atc cct att gtg ggc gtg att gaa cca agc 288  
 630 Met Gln Lys His Ser Lys Ile Pro Ile Val Gly Val Ile Glu Pro Ser  
 631 85 90 95  
 633 att tta gcg atc aag caa caa gta aaa gat aaa aac gcc tct att ttg 336  
 634 Ile Leu Ala Ile Lys Gln Gln Val Lys Asp Lys Asn Ala Ser Ile Leu  
 635 100 105 110  
 637 gtg cta ggg aca aaa gcg acg atc caa tcc aac gct tat gac aac gcc 384  
 638 Val Leu Gly Thr Lys Ala Thr Ile Gln Ser Asn Ala Tyr Asp Asn Ala  
 639 115 120 125  
 641 ctg aaa caa caa ggc tat ttg aat gtt tcg cat tta gcc act tct ctt 432  
 642 Leu Lys Gln Gln Gly Tyr Leu Asn Val Ser His Leu Ala Thr Ser Leu  
 643 130 135 140  
 645 ttt gtg cct ttg att gaa gaa agt att tta gag ggc gaa ttg cta gaa 480  
 646 Phe Val Pro Leu Ile Glu Glu Ser Ile Leu Glu Gly Glu Leu Leu Glu  
 647 145 150 155 160  
 649 act tgc atg cgt tat tat ttc act ccg tta gag atc ttg cct gaa gtg 528  
 650 Thr Cys Met Arg Tyr Tyr Phe Thr Pro Leu Glu Ile Leu Pro Glu Val  
 651 165 170 175  
 653 gtt att tta ggt tgc acg cat ttt ccc tta atc gct caa aaa att gag 576  
 654 Val Ile Leu Gly Cys Thr His Phe Pro Leu Ile Ala Gln Lys Ile Glu  
 655 180 185 190  
 657 ggc tat ttt atg gag cat ttt gcc ctt tca acg ccc ccc cta ctc atc 624  
 658 Gly Tyr Phe Met Glu His Phe Ala Leu Ser Thr Pro Pro Leu Leu Ile  
 659 195 200 205  
 661 cat tcg ggc gat gct att gtg gaa tat ttg cag caa aaa tac gcc ctt 672  
 662 His Ser Gly Asp Ala Ile Val Glu Tyr Leu Gln Gln Lys Tyr Ala Leu  
 663 210 215 220

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/729,571

DATE: 12/22/2003

TIME: 10:11:23

Input Set : A:\SEQUENCE LISTING.txt

Output Set: N:\CRF4\12222003\J729571.raw

665 aaa aaa aat gca cac gca ttc cct aaa gtg gaa ttt cat gcg agt ggc 720  
 666 Lys Lys Asn Ala His Ala Phe Pro Lys Val Glu Phe His Ala Ser Gly  
 667 225 230 235 240  
 669 gat gtg atc tgg cta gaa aaa cag gct aaa gaa tgg ctc aaa ttg taa 768  
 670 Asp Val Ile Trp Leu Glu Lys Gln Ala Lys Glu Trp Leu Lys Leu  
 671 245 250 255  
 743 <210> SEQ ID NO: 11  
 744 <211> LENGTH: 749  
 745 <212> TYPE: DNA  
 746 <213> ORGANISM: H. pylori ><220>  
 748 <221> NAME/KEY: CDS  
 749 <222> LOCATION: (1)..(747)  
 Ek-> 751 <400> SEQUENCE: 11  
 752 atg aaa ata ggc gtt ttt gat agc ggt gtg gga ggg ttt agc gtt tta 48  
 753 Met Lys Ile Gly Val Phe Asp Ser Gly Val Gly Gly Phe Ser Val Leu  
 754 1 5 10 15  
 756 aaa agc ctt tta aaa gcg caa att ttt gat gaa atc atc tat tat ggc 96  
 757 Lys Ser Leu Leu Lys Ala Gln Ile Phe Asp Glu Ile Ile Tyr Tyr Gly  
 758 20 25 30  
 760 gat agc gct aga gtg cct tat ggc act aaa gac ccc acc acg atc aag 144  
 761 Asp Ser Ala Arg Val Pro Tyr Gly Thr Lys Asp Pro Thr Thr Ile Lys  
 762 35 40 45  
 764 caa ttt ggc tta gag gct ttg gat ttt ttc aaa ccg cac cag att aaa 192  
 765 Gln Phe Gly Leu Glu Ala Leu Asp Phe Phe Lys Pro His Gln Ile Lys  
 766 50 55 60  
 768 tta ttg att gtg gca tgc aac aca gcg agc gct cta gct tta gaa gag 240  
 769 Leu Leu Ile Val Ala Cys Asn Thr Ala Ser Ala Leu Ala Leu Glu Glu  
 770 65 70 75 80  
 772 atg caa aag cat tcc aaa atc cct att gtg ggc gtg att gag cca agc 288  
 773 Met Gln Lys His Ser Lys Ile Pro Ile Val Gly Val Ile Glu Pro Ser  
 774 85 90 95  
 776 att tta gcg atc aag caa caa gta aaa gat aaa aac gcc cct att tta 336  
 777 Ile Leu Ala Ile Lys Gln Gln Val Lys Asp Lys Asn Ala Pro Ile Leu  
 778 100 105 110  
 780 gtg cta ggg aca aaa gcg acg att caa tct aac gct tat gac aac gcc 384  
 781 Val Leu Gly Thr Lys Ala Thr Ile Gln Ser Asn Ala Tyr Asp Asn Ala  
 782 115 120 125  
 784 cta aaa caa caa ggc tat ttg aac att tcg cat tta gcc act tct ctt 432  
 785 Leu Lys Gln Gln Gly Tyr Leu Asn Ile Ser His Leu Ala Thr Ser Leu  
 786 130 135 140  
 788 ttt gtg cct ttg att gaa gaa agt att tta gag ggc gaa ttg tta gag 480  
 789 Phe Val Pro Leu Ile Glu Glu Ser Ile Leu Glu Gly Glu Leu Leu Glu  
 790 145 150 155 160  
 792 act tgc atg cgt tat tat ttc act ccc tta aag att tta cct gaa gtg 528  
 793 Thr Cys Met Arg Tyr Tyr Phe Thr Pro Leu Lys Ile Leu Pro Glu Val  
 794 165 170 175  
 796 atc att tta ggt tgc acg cat ttt ccc ttg atc gct caa aaa att gag 576  
 797 Ile Ile Leu Gly Cys Thr His Phe Pro Leu Ile Ala Gln Lys Ile Glu  
 798 180 185 190



## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/729,571

DATE: 12/22/2003

TIME: 10:11:23

Input Set : A:\SEQUENCE LISTING.txt

Output Set : N:\CRF4\12222003\J729571.raw

```

800 ggc tat ttt atg gag cat ttt gcc ctt cca acc ccc ccc cta ctc atc      624
801 Gly Tyr Phe Met Glu His Phe Ala Leu Pro Thr Pro Pro Leu Leu Ile
802      195      200      205
804 cat tcg ggc gat gct att gta gaa tat ttg cag caa aaa tac acc ctt      672
805 His Ser Gly Asp Ala Ile Val Glu Tyr Leu Gln Gln Lys Tyr Thr Leu
806      210      215      220
808 aag aaa aat gca cac gca ttc cct aaa gtg gaa ttt cat gcg agt ggc      720
809 Lys Lys Asn Ala His Ala Phe Pro Lys Val Glu Phe His Ala Ser Gly
810 225      230      235      240
812 gat gtg gtt tgg cta gaa aaa cag gct aa      749
813 Asp Val Val Trp Leu Glu Lys Gln Ala
814      245
886 <210> SEQ ID NO: 13
887 <211> LENGTH: 768
888 <212> TYPE: DNA
889 <213> ORGANISM: H. pylori
891 <221> NAME/KEY: CDS
892 <222> LOCATION: (1)..(768)
894 <400> SEQUENCE: 13
895 atg aaa ata ggc gtt ttt gat agc ggt gtg gga ggg ttt agc gtt tta      48
896 Met Lys Ile Gly Val Phe Asp Ser Gly Val Gly Gly Phe Ser Val Leu
897 1      5      10      15
899 aaa agc ctt tta aaa gcg caa att ttt gat gaa atc atc tat tat ggc      96
900 Lys Ser Leu Leu Lys Ala Gln Ile Phe Asp Glu Ile Ile Tyr Tyr Gly
901      20      25      30
903 gat agt gct aga gtg cct tat ggc act aaa gac ccc acc acg atc aag      144
904 Asp Ser Ala Arg Val Pro Tyr Gly Thr Lys Asp Pro Thr Thr Ile Lys
905      35      40      45
907 caa ttt ggc tta gag gct ttg gat ttt ttc aaa ccg cac cag att gga      192
908 Gln Phe Gly Leu Glu Ala Leu Asp Phe Phe Lys Pro His Gln Ile Gly
909      50      55      60
911 tta ttg att gtg gca tgc aac aca gcg agc gct cta gct tta gaa gag      240
912 Leu Leu Ile Val Ala Cys Asn Thr Ala Ser Ala Leu Ala Leu Glu Glu
913 65      70      75      80
915 atg caa aag cat tcc aaa atc cct att gtg ggc gtg att gag cca agc      288
916 Met Gln Lys His Ser Lys Ile Pro Ile Val Gly Val Ile Glu Pro Ser
917      85      90      95
919 att tta gcg atc aaa caa caa gtg aaa gat aaa aac gct cct att tta      336
920 Ile Leu Ala Ile Lys Gln Gln Val Lys Asp Lys Asn Ala Pro Ile Leu
921      100      105      110
923 gtg cta ggg aca aaa gcg acg att caa tct aac gct tac gat aac gcc      384
924 Val Leu Gly Thr Lys Ala Thr Ile Gln Ser Asn Ala Tyr Asp Asn Ala
925      115      120      125
927 ctg aaa caa caa ggc tat ttg aat gtt tcg cat tta gcc act tct ctt      432
928 Leu Lys Gln Gln Gly Tyr Leu Asn Val Ser His Leu Ala Thr Ser Leu
929      130      135      140
931 ttt gtg cct ttg att gaa gaa aat att tta gag ggc gaa ttg cta gaa      480
932 Phe Val Pro Leu Ile Glu Glu Asn Ile Leu Glu Gly Glu Leu Leu Glu
933 145      150      155      160

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/729,571

DATE: 12/22/2003

TIME: 10:11:23

Input Set : A:\SEQUENCE LISTING.txt

Output Set: N:\CRF4\12222003\J729571.raw

```

935 act tgc atg cgt tat tat ttc act ccc tta aag att tta cct gaa gtg      528
936 Thr Cys Met Arg Tyr Tyr Phe Thr Pro Leu Lys Ile Leu Pro Glu Val
937          165          170          175
939 atc att tta ggt tgc acg cat ttt ccc ttg atc gct caa aaa att gag      576
940 Ile Ile Leu Gly Cys Thr His Phe Pro Leu Ile Ala Gln Lys Ile Glu
941          180          185          190
943 ggc tat ttt atg gag cat ttt gcc ctt tta acg ccc ccc cta ctc atc      624
944 Gly Tyr Phe Met Glu His Phe Ala Leu Leu Thr Pro Pro Leu Leu Ile
945          195          200          205
947 cat tct ggc gat gct att gta gaa tat ttg caa caa aaa tac gcc ctt      672
948 His Ser Gly Asp Ala Ile Val Glu Tyr Leu Gln Gln Lys Tyr Ala Leu
949          210          215          220
951 aag aaa aat gca cac tca ttc cct aaa gtg gaa ttt cat gcg agc ggc      720
952 Lys Lys Asn Ala His Ser Phe Pro Lys Val Glu Phe His Ala Ser Gly
953 225          230          235          240
955 gat gtg atc tgg cta gaa aaa cag gct aaa gaa tgg ctc aaa ttg taa      768
956 Asp Val Ile Trp Leu Glu Lys Gln Ala Lys Glu Trp Leu Lys Leu
957          245          250          255
1029 <210> SEQ ID NO: 15
1030 <211> LENGTH: 768
1031 <212> TYPE: DNA
1032 <213> ORGANISM: H. pylori
1034 <221> NAME/KEY: CDS
1035 <222> LOCATION: (1)..(768)
1037 <400> SEQUENCE: 15
1038 atg aaa ata ggc gtt ttt gat agc ggt gtg gga ggg ttt agc gtt tta      48
1039 Met Lys Ile Gly Val Phe Asp Ser Gly Val Gly Gly Phe Ser Val Leu
1040 1          5          10          15
1042 aaa agc ctt tta aaa gcg caa att ttt gat gaa atc atc tat tat ggc      96
1043 Lys Ser Leu Leu Lys Ala Gln Ile Phe Asp Glu Ile Ile Tyr Tyr Gly
1044          20          25          30
1046 gat agc gct aga gtg cct tat ggc act aaa gac ccc acc acg atc aag      144
1047 Asp Ser Ala Arg Val Pro Tyr Gly Thr Lys Asp Pro Thr Thr Ile Lys
1048          35          40          45
1050 caa ttt ggc tta gag gct ttg gat ttt ttc aaa ccg cac cag att gaa      192
1051 Gln Phe Gly Leu Glu Ala Leu Asp Phe Phe Lys Pro His Gln Ile Glu
1052          50          55          60
1054 tta ttg att gtg gca tgc aac aca gcg agc gct cta gct tta gaa gag      240
1055 Leu Leu Ile Val Ala Cys Asn Thr Ala Ser Ala Leu Ala Leu Glu Glu
1056 65          70          75          80
1058 atg caa aag cat tcc aaa atc cct att gtg ggc gtg att gaa cca agc      288
1059 Met Gln Lys His Ser Lys Ile Pro Ile Val Gly Val Ile Glu Pro Ser
1060          85          90          95
1062 att tta gcg atc aaa caa caa gtg aaa gat aaa aac gct cct att tta      336
1063 Ile Leu Ala Ile Lys Gln Gln Val Lys Asp Lys Asn Ala Pro Ile Leu
1064          100          105          110
1066 gtg cta ggg aca aaa gcg acg att caa tct aac gct tac gac aac gcc      384
1067 Val Leu Gly Thr Lys Ala Thr Ile Gln Ser Asn Ala Tyr Asp Asn Ala
1068          115          120          125

```

OK-&gt;

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/729,571

DATE: 12/22/2003

TIME: 10:11:23

Input Set : A:\SEQUENCE LISTING.txt

Output Set: N:\CRF4\12222003\J729571.raw

```

1070 ctg aaa caa caa ggc tat ttg aat gtt tcg cat tta gcc act tct ctt      432
1071 Leu Lys Gln Gln Gly Tyr Leu Asn Val Ser His Leu Ala Thr Ser Leu
1072      130                      135                      140
1074 ttt gtg cct ttg att gaa gaa agt att tta gag ggc gaa ttg cta gaa      480
1075 Phe Val Pro Leu Ile Glu Glu Ser Ile Leu Glu Gly Glu Leu Leu Glu
1076 145                      150                      155                      160
1078 act tgc atg cgt tat tat ttc act ccc tta aag att tta ccc aaa gta      528
1079 Thr Cys Met Arg Tyr Tyr Phe Thr Pro Leu Lys Ile Leu Pro Lys Val
1080                      165                      170                      175
1082 atc att tta ggt tgc acg cat ttt ccc ttg atc gct cac caa att aag      576
1083 Ile Ile Leu Gly Cys Thr His Phe Pro Leu Ile Ala His Gln Ile Lys
1084                      180                      185                      190
1086 ggc tat ttt atg ggg cat ttt gcc ctt tca acg ccc ccc cta ctc atc      624
1087 Gly Tyr Phe Met Gly His Phe Ala Leu Ser Thr Pro Pro Leu Leu Ile
1088      195                      200                      205
1090 cat tcg ggc gat gct att gtg gga tat ttg caa caa aaa tac gcc ctt      672
1091 His Ser Gly Asp Ala Ile Val Gly Tyr Leu Gln Gln Lys Tyr Ala Leu
1092      210                      215                      220
1094 aag aaa aat gca cac gca ttc cct aaa gtg gaa ttt cat gcg agc ggc      720
1095 Lys Lys Asn Ala His Ala Phe Pro Lys Val Glu Phe His Ala Ser Gly
1096 225                      230                      235                      240
1098 gat gtg atc tgg cta gaa aaa cag gct aaa gaa tgg ctc aaa ttg taa      768
1099 Asp Val Ile Trp Leu Glu Lys Gln Ala Lys Glu Trp Leu Lys Leu
1100                      245                      250                      255
1172 <210> SEQ ID NO: 17
1173 <211> LENGTH: 768
1174 <212> TYPE: DNA
1175 <213> ORGANISM: H. pylori
1177 <221> NAME/KEY: CDS
1178 <222> LOCATION: (1)..(768)
1180 <400> SEQUENCE: 17
1181 atg aaa ata ggc gtt ttt gat agc ggt gtg gga ggg ttt agc gtt tta      48
1182 Met Lys Ile Gly Val Phe Asp Ser Gly Val Gly Gly Phe Ser Val Leu
1183 1                      5                      10                      15
1185 aaa agc ctt tta aaa gcg caa tta ttt gat gaa atc atc tat tat ggc      96
1186 Lys Ser Leu Leu Lys Ala Gln Leu Phe Asp Glu Ile Ile Tyr Tyr Gly
1187      20                      25                      30
1189 gat agc gct aga gtg cct tat ggc act aaa gac ccc acc acg atc aag      144
1190 Asp Ser Ala Arg Val Pro Tyr Gly Thr Lys Asp Pro Thr Thr Ile Lys
1191      35                      40                      45
1193 caa ttt ggc tta gag gct ttg gat ttt ttc aaa ccg cac cag att aaa      192
1194 Gln Phe Gly Leu Glu Ala Leu Asp Phe Phe Lys Pro His Gln Ile Lys
1195      50                      55                      60
1197 tta ttg att gtg gca tgc aac aca gcg agt gct ctg gct tta gaa gag      240
1198 Leu Leu Ile Val Ala Cys Asn Thr Ala Ser Ala Leu Ala Leu Glu Glu
1199 65                      70                      75                      80
1201 atg caa aag cat tcc aaa atc cct att gtg ggc gtg att gag cca agc      288
1202 Met Gln Lys His Ser Lys Ile Pro Ile Val Gly Val Ile Glu Pro Ser
1203      85                      90                      95

```

EKG&gt;

→(2207

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/729,571

DATE: 12/22/2003

TIME: 10:11:23

Input Set : A:\SEQUENCE LISTING.txt

Output Set: N:\CRF4\12222003\J729571.raw

```

1205 att tta gcg atc aaa caa cag gta aaa gat aaa aac gcc ccc att tta      336
1206 Ile Leu Ala Ile Lys Gln Gln Val Lys Asp Lys Asn Ala Pro Ile Leu
1207          100          105          110
1209 gtg cta ggc aca aaa gcg acg att caa tct aac gct tac gat aac gct      384
1210 Val Leu Gly Thr Lys Ala Thr Ile Gln Ser Asn Ala Tyr Asp Asn Ala
1211          115          120          125
1213 ctg aaa cga caa ggc tat ttg aac gtt tcg cat tta gcc act tcc ctt      432
1214 Leu Lys Arg Gln Gly Tyr Leu Asn Val Ser His Leu Ala Thr Ser Leu
1215          130          135          140
1217 ttt gtg cct ttg att gaa gaa agt att tta gag ggc gaa ttg tta gaa      480
1218 Phe Val Pro Leu Ile Glu Glu Ser Ile Leu Glu Gly Glu Leu Leu Glu
1219 145          150          155          160
1221 act tgc atg cgt tat tat ttc act ccc tta aag att tta cct gaa gtg      528
1222 Thr Cys Met Arg Tyr Tyr Phe Thr Pro Leu Lys Ile Leu Pro Glu Val
1223          165          170          175
1225 atc att tta ggt tgt acg cat ttt ccc ttg atc gct caa aaa att gag      576
1226 Ile Ile Leu Gly Cys Thr His Phe Pro Leu Ile Ala Gln Lys Ile Glu
1227          180          185          190
1229 ggc tat ttt atg gaa cat ttt gcc ttt cca acg ccc ccc cta ctc atc      624
1230 Gly Tyr Phe Met Glu His Phe Ala Phe Pro Thr Pro Pro Leu Leu Ile
1231          195          200          205
1233 cat tcg ggc gat gct att gtg gaa tat ttg cag caa aaa tac gcc ctt      672
1234 His Ser Gly Asp Ala Ile Val Glu Tyr Leu Gln Gln Lys Tyr Ala Leu
1235          210          215          220
1237 aag aaa aat gca cac gca tta cct aaa gtg gaa ttt cat gcg agc ggc      720
1238 Lys Lys Asn Ala His Ala Leu Pro Lys Val Glu Phe His Ala Ser Gly
1239 225          230          235          240
1241 gat gtg atc tgg cta gaa aaa caa gct aaa gaa tgg ctc aaa ttg taa      768
1242 Asp Val Ile Trp Leu Glu Lys Gln Ala Lys Glu Trp Leu Lys Leu
1243          245          250          255
1315 <210> SEQ ID NO: 19
1316 <211> LENGTH: 768
1317 <212> TYPE: DNA
1318 <213> ORGANISM: H. pylori → (220)
1320 <221> NAME/KEY: CDS
1321 <222> LOCATION: (1)..(768)
1323 <400> SEQUENCE: 19
1324 atg aaa ata ggc gtt ttt gat agc ggt gtg gga ggg ttt agc gtt tta      48
1325 Met Lys Ile Gly Val Phe Asp Ser Gly Val Gly Gly Phe Ser Val Leu
1326 1          5          10          15
1328 aaa agc ctt tta aaa gcg caa tta ttt gat gaa atc atc tat tat ggc      96
1329 Lys Ser Leu Leu Lys Ala Gln Leu Phe Asp Glu Ile Ile Tyr Tyr Gly
1330          20          25          30
1332 gat agc gct aga gtg cct tat ggc act aaa gac ccc acc acg atc aag      144
1333 Asp Ser Ala Arg Val Pro Tyr Gly Thr Lys Asp Pro Thr Thr Ile Lys
1334          35          40          45
1336 caa ttt ggc tta gag gct ttg gat ttt ttc aaa ccg cac cag att aaa      192
1337 Gln Phe Gly Leu Glu Ala Leu Asp Phe Phe Lys Pro His Gln Ile Lys
1338          50          55          60

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/729,571

DATE: 12/22/2003

TIME: 10:11:23

Input Set : A:\SEQUENCE LISTING.txt

Output Set: N:\CRF4\12222003\J729571.raw

```

1340 tta ttg att gtg gca tgc aac aca gcg agc gct cta gct tta gaa gag      240
1341 Leu Leu Ile Val Ala Cys Asn Thr Ala Ser Ala Leu Ala Leu Glu Glu
1342 65              70              75              80
1344 atg caa aag cat tcc aaa atc cct att gtg ggc gtg att gag cca agc      288
1345 Met Gln Lys His Ser Lys Ile Pro Ile Val Gly Val Ile Glu Pro Ser
1346              85              90              95
1348 att tta gcg atc aaa caa caa gta aag gat aaa aac gcc ccc att tta      336
1349 Ile Leu Ala Ile Lys Gln Gln Val Lys Asp Lys Asn Ala Pro Ile Leu
1350              100              105              110
1352 gtg cta ggg aca aaa gcg acg att caa tct aac gct tac gat aac gct      384
1353 Val Leu Gly Thr Lys Ala Thr Ile Gln Ser Asn Ala Tyr Asp Asn Ala
1354              115              120              125
1356 ctg aaa caa caa ggc tat ttg aac gtt tcg cat tta gcc act tct ctt      432
1357 Leu Lys Gln Gln Gly Tyr Leu Asn Val Ser His Leu Ala Thr Ser Leu
1358              130              135              140
1360 ttt gtg cct ttg att gaa gaa aat att tta gag ggc gaa ttg tta gaa      480
1361 Phe Val Pro Leu Ile Glu Glu Asn Ile Leu Glu Gly Glu Leu Leu Glu
1362 145              150              155              160
1364 act tgc atg cgt tat tat ttc act ccc tta gag att tta cct gaa gtg      528
1365 Thr Cys Met Arg Tyr Tyr Phe Thr Pro Leu Glu Ile Leu Pro Glu Val
1366              165              170              175
1368 atc att tta ggt tgc acg cat ttt ccc tta atc gct caa aaa att gag      576
1369 Ile Ile Leu Gly Cys Thr His Phe Pro Leu Ile Ala Gln Lys Ile Glu
1370              180              185              190
1372 ggc tat ttc atg ggg cat ttt gcc ctt cca acg ccc ccc ata ctc atc      624
1373 Gly Tyr Phe Met Gly His Phe Ala Leu Pro Thr Pro Pro Ile Leu Ile
1374              195              200              205
1376 cat tct ggc gac gct att gta gaa tat ttg caa caa aaa tac gcc ctt      672
1377 His Ser Gly Asp Ala Ile Val Glu Tyr Leu Gln Gln Lys Tyr Ala Leu
1378              210              215              220
1380 aag aaa aat gca cac gca ttc cct aaa gtg gaa ttt cat gcg agc ggc      720
1381 Lys Lys Asn Ala His Ala Phe Pro Lys Val Glu Phe His Ala Ser Gly
1382 225              230              235              240
1384 gat atg atc tgg cta gaa aaa caa gct aaa gaa tgg ctc aaa ttg taa      768
1385 Asp Met Ile Trp Leu Glu Lys Gln Ala Lys Glu Trp Leu Lys Leu
1386              245              250              255
1458 <210> SEQ ID NO: 21
1459 <211> LENGTH: 768
1460 <212> TYPE: DNA
1461 <213> ORGANISM: H. pylori
1463 <221> NAME/KEY: CDS
1464 <222> LOCATION: (1)..(768)
1466 <400> SEQUENCE: 21
1467 atg aaa ata ggc gtt ttt gat agc ggt gtg gga ggg ttt agc gtt tta      48
1468 Met Lys Ile Gly Val Phe Asp Ser Gly Val Gly Gly Phe Ser Val Leu
1469 1              5              10              15
1471 aaa agc ctt tta aaa gcg caa tta ttt gat gaa atc atc tat tat ggc      96
1472 Lys Ser Leu Leu Lys Ala Gln Leu Phe Asp Glu Ile Ile Tyr Tyr Gly
1473              20              25              30

```

1466 <400> SEQUENCE: 21

1467 atg aaa ata ggc gtt ttt gat agc ggt gtg gga ggg ttt agc gtt tta 48  
 1468 Met Lys Ile Gly Val Phe Asp Ser Gly Val Gly Gly Phe Ser Val Leu  
 1469 1 5 10 15  
 1471 aaa agc ctt tta aaa gcg caa tta ttt gat gaa atc atc tat tat ggc 96  
 1472 Lys Ser Leu Leu Lys Ala Gln Leu Phe Asp Glu Ile Ile Tyr Tyr Gly  
 1473 20 25 30

## RAW SEQUENCE LISTING

DATE: 12/22/2003

PATENT APPLICATION: US/10/729,571

TIME: 10:11:23

Input Set : A:\SEQUENCE LISTING.txt

Output Set: N:\CRF4\12222003\J729571.raw

```

1475 gat agc gct aga gtg cct tat ggc act aaa gac ccc acc acg atc aag      144
1476 Asp Ser Ala Arg Val Pro Tyr Gly Thr Lys Asp Pro Thr Thr Ile Lys
1477          35          40          45
1479 caa ttt ggc tta gag gct ttg gat ttt ttc aaa ccg cac cag att aaa      192
1480 Gln Phe Gly Leu Glu Ala Leu Asp Phe Phe Lys Pro His Gln Ile Lys
1481          50          55          60
1483 tta ttg att gta gca tgc aac aca gcg agc gct cta gct tta gaa gag      240
1484 Leu Leu Ile Val Ala Cys Asn Thr Ala Ser Ala Leu Ala Leu Glu Glu
1485 65          70          75          80
1487 atg caa aag cat tcc aaa atc cct att gtg ggc gtg att gag cca agc      288
1488 Met Gln Lys His Ser Lys Ile Pro Ile Val Gly Val Ile Glu Pro Ser
1489          85          90          95
1491 att tta gcg atc aaa caa caa gta aaa gat aaa aac gcc cct att tta      336
1492 Ile Leu Ala Ile Lys Gln Gln Val Lys Asp Lys Asn Ala Pro Ile Leu
1493          100          105          110
1495 gtg cta ggg aca aaa gcg acg att caa tct aac gct tat gac aac gcc      384
1496 Val Leu Gly Thr Lys Ala Thr Ile Gln Ser Asn Ala Tyr Asp Asn Ala
1497          115          120          125
1499 ctg aaa caa caa ggc tat ttg aat gtt tcg cat tta gcc act tct ctt      432
1500 Leu Lys Gln Gln Gly Tyr Leu Asn Val Ser His Leu Ala Thr Ser Leu
1501          130          135          140
1503 ttt gtg cct ttg att gaa gaa agt att tta gag ggc gaa ttg tta gaa      480
1504 Phe Val Pro Leu Ile Glu Glu Ser Ile Leu Glu Gly Glu Leu Leu Glu
1505 145          150          155          160
1507 act tgc atg cgt tat tat ttc act ccc tta aag att tta cct gaa gtg      528
1508 Thr Cys Met Arg Tyr Tyr Phe Thr Pro Leu Lys Ile Leu Pro Glu Val
1509          165          170          175
1511 att att tta ggt tgc acg cat ttt ccc ttg atc gct caa aaa att gag      576
1512 Ile Ile Leu Gly Cys Thr His Phe Pro Leu Ile Ala Gln Lys Ile Glu
1513          180          185          190
1515 agc tat ttt atg ggg cat ttt gcc ctt cca acg ccc ccc cta ctc atc      624
1516 Ser Tyr Phe Met Gly His Phe Ala Leu Pro Thr Pro Pro Leu Leu Ile
1517          195          200          205
1519 cat tct ggc gat gct att gtg gaa tat ttg cag caa aaa tac gcc ctt      672
1520 His Ser Gly Asp Ala Ile Val Glu Tyr Leu Gln Gln Lys Tyr Ala Leu
1521          210          215          220
1523 aag aaa aac gca cac gca ttc cct aaa gtg gaa ttt cat gcg agc ggc      720
1524 Lys Lys Asn Ala His Ala Phe Pro Lys Val Glu Phe His Ala Ser Gly
1525 225          230          235          240
1527 gat gtg atc tgg cta gaa aaa caa gct aaa gaa tgg ctc aaa ttg taa      768
1528 Asp Val Ile Trp Leu Glu Lys Gln Ala Lys Glu Trp Leu Lys Leu
1529          245          250          255

```

1601 &lt;210&gt; SEQ ID NO: 23


1602 &lt;211&gt; LENGTH: 768

1603 &lt;212&gt; TYPE: DNA

1604 <213> ORGANISM: H. pylori }<2207

1607 &lt;221&gt; NAME/KEY: CDS

1608 &lt;222&gt; LOCATION: (1)..(768)

 1610 <400> SEQUENCE: 23

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/729,571

DATE: 12/22/2003

TIME: 10:11:23

Input Set : A:\SEQUENCE LISTING.txt

Output Set: N:\CRF4\12222003\J729571.raw

```

1611 atg aaa ata ggc gtt ttt gat agc ggt gtg gga ggg ttt agc gtt tta      48
1612 Met Lys Ile Gly Val Phe Asp Ser Gly Val Gly Gly Phe Ser Val Leu
1613 1          5          10          15
1615 aaa agc ctt tta aaa gcg caa cta ttt gat gaa atc atc tat tat ggc      96
1616 Lys Ser Leu Leu Lys Ala Gln Leu Phe Asp Glu Ile Ile Tyr Tyr Gly
1617          20          25          30
1619 gat agc gct aga gtg cct tat ggc act aaa gac ccc acc acg atc aag      144
1620 Asp Ser Ala Arg Val Pro Tyr Gly Thr Lys Asp Pro Thr Thr Ile Lys
1621          35          40          45
1623 caa ttt ggc tta gag gct ttg gat ttt ttc aaa ccg cac cag att gga      192
1624 Gln Phe Gly Leu Glu Ala Leu Asp Phe Phe Lys Pro His Gln Ile Gly
1625          50          55          60
1627 tta ttg att gtg gca tgc aac aca gcg agc gct ctg gct tta gaa gag      240
1628 Leu Leu Ile Val Ala Cys Asn Thr Ala Ser Ala Leu Ala Leu Glu Glu
1629 65          70          75          80
1631 atg caa aaa tat tcc aaa atc cct att gtg ggc gtg att gag cca agc      288
1632 Met Gln Lys Tyr Ser Lys Ile Pro Ile Val Gly Val Ile Glu Pro Ser
1633          85          90          95
1635 att tta gcg atc aaa caa caa gta aaa gat aaa aac gcc ccc att tta      336
1636 Ile Leu Ala Ile Lys Gln Gln Val Lys Asp Lys Asn Ala Pro Ile Leu
1637          100          105          110
1639 gtg cta ggg aca aaa gcg acg atc caa tct aac gct tat gat aac gcc      384
1640 Val Leu Gly Thr Lys Ala Thr Ile Gln Ser Asn Ala Tyr Asp Asn Ala
1641          115          120          125
1643 ctg aaa caa caa ggc tat ttg aac att tcg cat tta gcc act tct ctt      432
1644 Leu Lys Gln Gln Gly Tyr Leu Asn Ile Ser His Leu Ala Thr Ser Leu
1645          130          135          140
1647 ttt gtg ccc ttg att gaa gaa agt att tta gag ggc gaa ttg tta gaa      480
1648 Phe Val Pro Leu Ile Glu Glu Ser Ile Leu Glu Gly Glu Leu Leu Glu
1649 145          150          155          160
1651 act tgc atg cgt tat tat ttc act cca tta gag att tta cct gaa gtg      528
1652 Thr Cys Met Arg Tyr Tyr Phe Thr Pro Leu Glu Ile Leu Pro Glu Val
1653          165          170          175
1655 atc att tta ggt tgc acg cat ttt ccc ttg atc gct caa aaa att gag      576
1656 Ile Ile Leu Gly Cys Thr His Phe Pro Leu Ile Ala Gln Lys Ile Glu
1657          180          185          190
1659 agc tat ttt atg gag cat ttt gcc ctt tca acg ccc ccc tta ctc atc      624
1660 Ser Tyr Phe Met Glu His Phe Ala Leu Ser Thr Pro Pro Leu Leu Ile
1661          195          200          205
1663 cat tct ggc gat gct att gtg gaa tac ttg caa caa aaa tac gcc ctt      672
1664 His Ser Gly Asp Ala Ile Val Glu Tyr Leu Gln Gln Lys Tyr Ala Leu
1665          210          215          220
1667 aag aaa aac gca cac gca ttc cct aaa gtg gaa ttt cat gcg agc ggc      720
1668 Lys Lys Asn Ala His Ala Phe Pro Lys Val Glu Phe His Ala Ser Gly
1669 225          230          235          240
1671 gat gtg atc tgg cta gaa aaa cag gct aaa gaa tgg ctc aaa ttg taa      768
1672 Asp Val Ile Trp Leu Glu Lys Gln Ala Lys Glu Trp Leu Lys Leu
1673          245          250          255
1745 <210> SEQ ID NO: 25

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/729,571

DATE: 12/22/2003

TIME: 10:11:23

Input Set : A:\SEQUENCE LISTING.txt

Output Set: N:\CRF4\12222003\J729571.raw

1746 &lt;211&gt; LENGTH: 768

1747 &lt;212&gt; TYPE: DNA

1748 <213> ORGANISM: *H. pylori*

1750 &lt;221&gt; NAME/KEY: CDS

1751 &lt;222&gt; LOCATION: (1)..(768)

OK-&gt; 1753 &lt;400&gt; SEQUENCE: 25

1754 atg aaa ata ggc gtt ttt gat agc ggt gtg gga ggg ttt agc gtt tta 48  
 1755 Met Lys Ile Gly Val Phe Asp Ser Gly Val Gly Gly Phe Ser Val Leu  
 1756 1 5 10 15  
 1758 aaa agc ctt tta aaa gcg caa tta ttt gat gaa atc atc tat tat ggc 96  
 1759 Lys Ser Leu Leu Lys Ala Gln Leu Phe Asp Glu Ile Ile Tyr Tyr Gly  
 1760 20 25 30  
 1762 gat agc gct aga gtg cct tat ggc act aaa gac ccc acc acg atc aag 144  
 1763 Asp Ser Ala Arg Val Pro Tyr Gly Thr Lys Asp Pro Thr Thr Ile Lys  
 1764 35 40 45  
 1766 caa ttt ggc tta gag gct ttg gat ttt ttc aaa ccg cac aaa att gaa 192  
 1767 Gln Phe Gly Leu Glu Ala Leu Asp Phe Phe Lys Pro His Lys Ile Glu  
 1768 50 55 60  
 1770 tta tta att gtg gca tgc aac aca gcg agc gct ctg gct tta gaa gag 240  
 1771 Leu Leu Ile Val Ala Cys Asn Thr Ala Ser Ala Leu Ala Leu Glu Glu  
 1772 65 70 75 80  
 1774 atg caa aag cat tcc aaa atc ccc att gtg ggc gtg att gag cca agc 288  
 1775 Met Gln Lys His Ser Lys Ile Pro Ile Val Gly Val Ile Glu Pro Ser  
 1776 85 90 95  
 1778 att tta gcg atc aaa caa caa gtg aaa gat aaa aac acc cct att tta 336  
 1779 Ile Leu Ala Ile Lys Gln Gln Val Lys Asp Lys Asn Thr Pro Ile Leu  
 1780 100 105 110  
 1782 gtg cta ggg aca aaa gcg acg atc caa tct aac gct tac gat aac gcc 384  
 1783 Val Leu Gly Thr Lys Ala Thr Ile Gln Ser Asn Ala Tyr Asp Asn Ala  
 1784 115 120 125  
 1786 ctg aaa caa caa ggc tat ttg aag gtt tcg cat ttg gcc act tct ctt 432  
 1787 Leu Lys Gln Gln Gly Tyr Leu Lys Val Ser His Leu Ala Thr Ser Leu  
 1788 130 135 140  
 1790 ttt gtg cct ttg att gaa gaa agt att tta gag ggc gaa ttg tta gaa 480  
 1791 Phe Val Pro Leu Ile Glu Glu Ser Ile Leu Glu Gly Glu Leu Leu Glu  
 1792 145 150 155 160  
 1794 act tgc atg cgt tat tat ttc act cca tta gaa atc tta cct gaa gtg 528  
 1795 Thr Cys Met Arg Tyr Tyr Phe Thr Pro Leu Glu Ile Leu Pro Glu Val  
 1796 165 170 175  
 1798 gtt att tta ggc tgc acg cat ttt ccc ttg atc gct caa aaa att gag 576  
 1799 Val Ile Leu Gly Cys Thr His Phe Pro Leu Ile Ala Gln Lys Ile Glu  
 1800 180 185 190  
 1802 ggc tat ttt atg gaa cat ttt gcc ctt cca acg ccc ccc cta ctc atc 624  
 1803 Gly Tyr Phe Met Glu His Phe Ala Leu Pro Thr Pro Pro Leu Leu Ile  
 1804 195 200 205  
 1806 cat tct ggc gac gct att gtg gga tat ttg cag caa aaa tac gcc ctt 672  
 1807 His Ser Gly Asp Ala Ile Val Gly Tyr Leu Gln Gln Lys Tyr Ala Leu  
 1808 210 215 220  
 1810 aag aaa aac gca cac gca ttc cct aaa gtg gaa ttt cat gcg agc ggc 720



## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/729,571

DATE: 12/22/2003

TIME: 10:11:23

Input Set : A:\SEQUENCE LISTING.txt

Output Set : N:\CRF4\12222003\J729571.raw

```

1811 Lys Lys Asn Ala His Ala Phe Pro Lys Val Glu Phe His Ala Ser Gly
1812 225                230                235                240
1814 gat gta att tgg cta gaa aaa cag gct aaa gaa tgg ctc aaa ttg taa      768
1815 Asp Val Ile Trp Leu Glu Lys Gln Ala Lys Glu Trp Leu Lys Leu
1816                245                250                255
1888 <210> SEQ ID NO: 27
1889 <211> LENGTH: 768
1890 <212> TYPE: DNA
1891 <213> ORGANISM: H. pylori
1893 <221> NAME/KEY: CDS
1894 <222> LOCATION: (1)..(768)
OK-> 1896 <400> SEQUENCE: 27
1897 atg aaa ata ggc gtt ttt gat agc ggt gtg gga ggg ttt agc gtt tta      48
1898 Met Lys Ile Gly Val Phe Asp Ser Gly Val Gly Gly Phe Ser Val Leu
1899 1                5                10                15
1901 aaa agc ctt tta aaa gcg caa att ttt gat gaa atc atc tat tat ggc      96
1902 Lys Ser Leu Leu Lys Ala Gln Ile Phe Asp Glu Ile Ile Tyr Tyr Gly
1903                20                25                30
1905 gat agc gct agg gtg cct tat ggc act aaa gac ccc acc acg atc aag      144
1906 Asp Ser Ala Arg Val Pro Tyr Gly Thr Lys Asp Pro Thr Thr Ile Lys
1907                35                40                45
1909 caa ttt ggc tta gag gct ttg gat ttt ttc aaa ccg cac aag att gaa      192
1910 Gln Phe Gly Leu Glu Ala Leu Asp Phe Phe Lys Pro His Lys Ile Glu
1911                50                55                60
1913 tta ttg att gtg gca tgc aac aca gcg agc gct cta gct tta gaa gaa      240
1914 Leu Leu Ile Val Ala Cys Asn Thr Ala Ser Ala Leu Ala Leu Glu Glu
1915 65                70                75                80
1917 atg caa aag cat tcc aaa atc cct att gtg ggc gtg att gaa cca agc      288
1918 Met Gln Lys His Ser Lys Ile Pro Ile Val Gly Val Ile Glu Pro Ser
1919                85                90                95
1921 att tta gcg atc aaa caa caa gta aaa gat aaa aac gcc cct att tta      336
1922 Ile Leu Ala Ile Lys Gln Gln Val Lys Asp Lys Asn Ala Pro Ile Leu
1923                100               105               110
1925 gtg cta ggg aca aaa gcg acg att caa tct aac gct tat gac aac gcc      384
1926 Val Leu Gly Thr Lys Ala Thr Ile Gln Ser Asn Ala Tyr Asp Asn Ala
1927                115               120               125
1929 ctg aaa caa caa ggc tat ttg aat gtt tcg cat tta gcc act tct ott      432
1930 Leu Lys Gln Gln Gly Tyr Leu Asn Val Ser His Leu Ala Thr Ser Leu
1931                130               135               140
1933 ttt gtg cct ttg att gaa gaa aat att tta gag ggc gaa ttg cta gaa      480
1934 Phe Val Pro Leu Ile Glu Glu Asn Ile Leu Glu Gly Glu Leu Leu Glu
1935 145               150               155               160
1937 act tgc atg cgt tat tat ttc act cca tta gag atc ttg cct gaa gtg      528
1938 Thr Cys Met Arg Tyr Tyr Phe Thr Pro Leu Glu Ile Leu Pro Glu Val
1939                165               170               175
1941 gtt att tta ggc tgc acg cat ttt ccc ttg atc gct cac caa att gag      576
1942 Val Ile Leu Gly Cys Thr His Phe Pro Leu Ile Ala His Gln Ile Glu
1943                180               185               190
1945 ggc tat ttt atg gag cat ttt gcc ctt tca acg ccc ccc cta ctc atc      624

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/729,571

DATE: 12/22/2003

TIME: 10:11:23

Input Set : A:\SEQUENCE LISTING.txt

Output Set: N:\CRF4\12222003\J729571.raw

```

1946 Gly Tyr Phe Met Glu His Phe Ala Leu Ser Thr Pro Pro Leu Leu Ile
1947          195          200          205
1949 cat tct ggc gat gct att gtg gaa tat ttg cag caa aaa tac gcc ctt      672
1950 His Ser Gly Asp Ala Ile Val Glu Tyr Leu Gln Gln Lys Tyr Ala Leu
1951      210          215          220
1953 aag aaa aac gca tgt gca ttc cct aaa gta gaa ttt cat gcg agc ggc      720
1954 Lys Lys Asn Ala Cys Ala Phe Pro Lys Val Glu Phe His Ala Ser Gly
1955 225          230          235          240
1957 gat gta att tgg cta gaa aaa cag gct aaa gaa tgg ctc aaa ttg taa      768
1958 Asp Val Ile Trp Leu Glu Lys Gln Ala Lys Glu Trp Leu Lys Leu
1959          245          250          255
2031 <210> SEQ ID NO: 29
2032 <211> LENGTH: 768
2033 <212> TYPE: DNA
2034 <213> ORGANISM: H. pylori
2036 <221> NAME/KEY: CDS
2037 <222> LOCATION: (1)..(768)
2039 <400> SEQUENCE: 29
2040 atg aaa ata ggc gtt ttt gat agc ggt gtg gga ggg ttt agc gtt tta      48
2041 Met Lys Ile Gly Val Phe Asp Ser Gly Val Gly Gly Phe Ser Val Leu
2042 1          5          10          15
2044 aaa agc ctt tta aaa gtg caa tta ttt gat gaa atc atc tat tat ggc      96
2045 Lys Ser Leu Leu Lys Val Gln Leu Phe Asp Glu Ile Ile Tyr Tyr Gly
2046          20          25          30
2048 gat agt gct agg gtg cct tat ggc act aaa gac ccc acc acg atc aag      144
2049 Asp Ser Ala Arg Val Pro Tyr Gly Thr Lys Asp Pro Thr Thr Ile Lys
2050          35          40          45
2052 caa ttt ggc tta gag gct ttg gat ttt ttc aaa ccg cac aag att gaa      192
2053 Gln Phe Gly Leu Glu Ala Leu Asp Phe Phe Lys Pro His Lys Ile Glu
2054          50          55          60
2056 tta ttg att gtg gca tgc aac aca gcg agc gct cta gct tta gga gag      240
2057 Leu Leu Ile Val Ala Cys Asn Thr Ala Ser Ala Leu Ala Leu Gly Glu
2058 65          70          75          80
2060 atg caa aag tat tcc aaa atc cct att gtg ggc gtg att gag oca agc      288
2061 Met Gln Lys Tyr Ser Lys Ile Pro Ile Val Gly Val Ile Glu Pro Ser
2062          85          90          95
2064 att tta gcg atc aaa caa caa gta aaa gat aaa aac gcc cct att tta      336
2065 Ile Leu Ala Ile Lys Gln Gln Val Lys Asp Lys Asn Ala Pro Ile Leu
2066          100          105          110
2068 gta cta ggg aca aaa gcg acg att cga tcc aac gct tat gac aac gcc      384
2069 Val Leu Gly Thr Lys Ala Thr Ile Arg Ser Asn Ala Tyr Asp Asn Ala
2070          115          120          125
2072 ctg aaa caa caa ggc tat ttg aat att tcg cat tta gcc act tct ctt      432
2073 Leu Lys Gln Gln Gly Tyr Leu Asn Ile Ser His Leu Ala Thr Ser Leu
2074          130          135          140
2076 ttt gtg cct ttg att gaa gaa aat att tta gag ggc gaa ttg cta gaa      480
2077 Phe Val Pro Leu Ile Glu Glu Asn Ile Leu Glu Gly Glu Leu Leu Glu
2078 145          150          155          160
2080 act tgc atg cgt tat tat ttc act cca tta gag att tta cct gaa gtg      528

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/729,571

DATE: 12/22/2003

TIME: 10:11:23

Input Set : A:\SEQUENCE LISTING.txt

Output Set: N:\CRF4\12222003\J729571.raw

```

2081 Thr Cys Met Arg Tyr Tyr Phe Thr Pro Leu Glu Ile Leu Pro Glu Val
2082          165          170          175
2084 gtt att tta ggt tgc acg cat ttt ccc ttg atc gct cac caa att gag      576
2085 Val Ile Leu Gly Cys Thr His Phe Pro Leu Ile Ala His Gln Ile Glu
2086          180          185          190
2088 ggc tat ttt atg gag cat ttt gcc ctt tca acg ccc ccc cta ctc atc      624
2089 Gly Tyr Phe Met Glu His Phe Ala Leu Ser Thr Pro Pro Leu Leu Ile
2090          195          200          205
2092 cat tct ggc gat gct att gtg gaa tat ttg caa caa aaa tac gcc ctt      672
2093 His Ser Gly Asp Ala Ile Val Glu Tyr Leu Gln Gln Lys Tyr Ala Leu
2094          210          215          220
2096 aag aaa aac gca tgc gca ttc cct aaa gta gaa ttc cat gcg agc ggc      720
2097 Lys Lys Asn Ala Cys Ala Phe Pro Lys Val Glu Phe His Ala Ser Gly
2098 225          230          235          240
2100 gat gta att tgg cta gaa aaa cag gct aaa gaa tgg ctc aaa ttg taa      768
2101 Asp Val Ile Trp Leu Glu Lys Gln Ala Lys Glu Trp Leu Lys Leu
2102          245          250          255
2174 <210> SEQ ID NO: 31
2175 <211> LENGTH: 768
2176 <212> TYPE: DNA
2177 <213> ORGANISM: H. pylori
2179 <221> NAME/KEY: CDS
2180 <222> LOCATION: (1)..(768)
@-> 2182 <400> SEQUENCE: 31
2183 atg aaa ata ggc gtt ttt gat agc ggt gtg gga ggg ttt agc gtt tta      48
2184 Met Lys Ile Gly Val Phe Asp Ser Gly Val Gly Gly Phe Ser Val Leu
2185 1          5          10          15
2187 aaa agc ctt tta aaa gcg caa att ttt gat gaa atc atc tat tat ggc      96
2188 Lys Ser Leu Leu Lys Ala Gln Ile Phe Asp Glu Ile Ile Tyr Tyr Gly
2189          20          25          30
2191 gat agt gct aga gtg cct tat ggc act aaa gac ccc acc acg atc aag      144
2192 Asp Ser Ala Arg Val Pro Tyr Gly Thr Lys Asp Pro Thr Thr Ile Lys
2193          35          40          45
2195 caa ttt ggc tta gag gct ttg gat ttt ttc aaa ccg cac cag att gga      192
2196 Gln Phe Gly Leu Glu Ala Leu Asp Phe Phe Lys Pro His Gln Ile Gly
2197          50          55          60
2199 tta ttg att gtg gca tgc aac aca gcg agc gct cta gct tta gaa gag      240
2200 Leu Leu Ile Val Ala Cys Asn Thr Ala Ser Ala Leu Ala Leu Glu Glu
2201 65          70          75          80
2203 atg caa aag cat tcc aaa atc cct att gtg ggt gtg att gag cca agc      288
2204 Met Gln Lys His Ser Lys Ile Pro Ile Val Gly Val Ile Glu Pro Ser
2205          85          90          95
2207 att tta gcg atc aaa caa caa gta aaa gat aaa aac gcc cct att tta      336
2208 Ile Leu Ala Ile Lys Gln Gln Val Lys Asp Lys Asn Ala Pro Ile Leu
2209          100          105          110
2211 gtg tta ggg aca aaa gcg acg att caa tcc aac gct tat gac aac gcc      384
2212 Val Leu Gly Thr Lys Ala Thr Ile Gln Ser Asn Ala Tyr Asp Asn Ala
2213          115          120          125
2215 ctg aaa caa caa ggc tat ttg aac gtt tcg cat tta gcc act tct ctt      432

```

## RAW SEQUENCE LISTING

DATE: 12/22/2003

PATENT APPLICATION: US/10/729,571

TIME: 10:11:23

Input Set : A:\SEQUENCE LISTING.txt

Output Set: N:\CRF4\12222003\J729571.raw

2216 Leu Lys Gln Gln Gly Tyr Leu Asn Val Ser His Leu Ala Thr Ser Leu  
 2217 130 135 140  
 2219 ttt gtg cct ttg att gaa gaa aat att tta gag ggc gaa ttg tta gaa 480  
 2220 Phe Val Pro Leu Ile Glu Glu Asn Ile Leu Glu Gly Glu Leu Leu Glu  
 2221 145 150 155 160  
 2223 act tgc atg cgt tat tat ttc act cca tta gag att tta cct gaa gtg 528  
 2224 Thr Cys Met Arg Tyr Tyr Phe Thr Pro Leu Glu Ile Leu Pro Glu Val  
 2225 165 170 175  
 2227 gtt att tta ggt tgc acg cat ttt ccc ttg atc gct cac caa att gag 576  
 2228 Val Ile Leu Gly Cys Thr His Phe Pro Leu Ile Ala His Gln Ile Glu  
 2229 180 185 190  
 2231 ggc tat ttt atg gag cat ttt gcc ctt tca acg ccc ccc tta ctc atc 624  
 2232 Gly Tyr Phe Met Glu His Phe Ala Leu Ser Thr Pro Pro Leu Leu Ile  
 2233 195 200 205  
 2235 cat tct ggc gat gct att gtg gaa tat ttg caa caa aaa tac acc ctt 672  
 2236 His Ser Gly Asp Ala Ile Val Glu Tyr Leu Gln Gln Lys Tyr Thr Leu  
 2237 210 215 220  
 2239 aag aaa aat gca tgc gcg ttc cct aaa gtg gaa ttt cat gcg agc ggc 720  
 2240 Lys Lys Asn Ala Cys Ala Phe Pro Lys Val Glu Phe His Ala Ser Gly  
 2241 225 230 235 240  
 2243 gat gtg gtt tgg cta gaa aaa cag gct aaa gaa tgg ctc aaa ttg taa 768  
 2244 Asp Val Val Trp Leu Glu Lys Gln Ala Lys Glu Trp Leu Lys Leu  
 2245 245 250 255  
 2317 <210> SEQ ID NO: 33  
 2318 <211> LENGTH: 765  
 2319 <212> TYPE: DNA  
 2320 <213> ORGANISM: H. pylori  
 2322 <221> NAME/KEY: CDS  
 2323 <222> LOCATION: (1)..(765)  
 2325 <400> SEQUENCE: 33  
 2326 atg aaa ata ggc gtt ttt gat agc ggt gtg gga ggg ttt agc gtt tta 48  
 2327 Met Lys Ile Gly Val Phe Asp Ser Gly Val Gly Gly Phe Ser Val Leu  
 2328 1 5 10 15  
 2330 aaa agc ctt tta aaa gcg caa cta ttt gat gaa atc atc tat tat ggc 96  
 2331 Lys Ser Leu Leu Lys Ala Gln Leu Phe Asp Glu Ile Ile Tyr Tyr Gly  
 2332 20 25 30  
 2334 gat agc gct aga gtg cct tat ggc act aaa gac ccc acc acg atc aag 144  
 2335 Asp Ser Ala Arg Val Pro Tyr Gly Thr Lys Asp Pro Thr Thr Ile Lys  
 2336 35 40 45  
 2338 caa ttt ggc tta gag gct ttg gat ttt ttc aaa ccg cac cag att aaa 192  
 2339 Gln Phe Gly Leu Glu Ala Leu Asp Phe Phe Lys Pro His Gln Ile Lys  
 2340 50 55 60  
 2342 tta ttg att gtg gca tgc aac acc gca agc gct ctg gct tta gaa gag 240  
 2343 Leu Leu Ile Val Ala Cys Asn Thr Ala Ser Ala Leu Ala Leu Glu Glu  
 2344 65 70 75 80  
 2346 atg caa aag cat tcc aaa atc cct gtt gtg ggc gtg att gag cca agc 288  
 2347 Met Gln Lys His Ser Lys Ile Pro Val Val Gly Val Ile Glu Pro Ser  
 2348 85 90 95  
 2350 att tta gcg atc aaa cgg caa gtg aaa gat aaa aac gcc cct att ttg 336

*see p.21 for more errors*

OK->

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/729,571

DATE: 12/22/2003

TIME: 10:11:23

Input Set : A:\SEQUENCE LISTING.txt

Output Set: N:\CRF4\12222003\J729571.raw

2351	Ile	Leu	Ala	Ile	Lys	Arg	Gln	Val	Lys	Asp	Lys	Asn	Ala	Pro	Ile	Leu	
2352				100					105					110			
2354	gtg	cta	ggg	aca	aaa	gcg	acg	att	caa	tcc	aac	gcc	tat	gat	aac	gcc	384
2355	Val	Leu	Gly	Thr	Lys	Ala	Thr	Ile	Gln	Ser	Asn	Ala	Tyr	Asp	Asn	Ala	
2356			115					120					125				
2358	ctg	aaa	caa	caa	ggc	tat	ttg	aat	gtt	tcg	cat	tta	gcc	act	tct	ctt	432
2359	Leu	Lys	Gln	Gln	Gly	Tyr	Leu	Asn	Val	Ser	His	Leu	Ala	Thr	Ser	Leu	
2360			130					135					140				
2362	ttt	gtg	cct	ttg	att	gaa	gaa	agt	att	tta	gag	ggc	gaa	ttg	cta	gaa	480
2363	Phe	Val	Pro	Leu	Ile	Glu	Glu	Ser	Ile	Leu	Glu	Gly	Glu	Leu	Leu	Glu	
2364	145					150					155					160	
2366	act	tgc	atg	cgt	tat	tat	ttc	act	cca	tta	gag	att	tta	cct	gaa	gtg	528
2367	Thr	Cys	Met	Arg	Tyr	Tyr	Phe	Thr	Pro	Leu	Glu	Ile	Leu	Pro	Glu	Val	
2368				165						170					175		
2370	gtt	att	tta	ggt	tgc	acg	cat	ttt	ccc	ttg	atc	gct	caa	aaa	att	gag	576
2371	Val	Ile	Leu	Gly	Cys	Thr	His	Phe	Pro	Leu	Ile	Ala	Gln	Lys	Ile	Glu	
2372			180						185					190			
2374	ggc	tat	ttt	atg	gag	cat	ttt	gcc	ctt	tca	acg	ccc	ccc	cta	ctc	atc	624
2375	Gly	Tyr	Phe	Met	Glu	His	Phe	Ala	Leu	Ser	Thr	Pro	Pro	Leu	Leu	Ile	
2376			195					200					205				
2378	cat	tct	ggc	gat	gct	att	gtg	gaa	tat	ttg	caa	caa	aat	tac	gcc	ctt	672
2379	His	Ser	Gly	Asp	Ala	Ile	Val	Glu	Tyr	Leu	Gln	Gln	Asn	Tyr	Ala	Leu	
2380			210					215					220				
2382	aag	aaa	aac	gca	tgc	gcg	ttc	cct	aaa	gtg	gaa	ttt	cat	gcg	agc	ggc	720
2383	Lys	Lys	Asn	Ala	Cys	Ala	Phe	Pro	Lys	Val	Glu	Phe	His	Ala	Ser	Gly	
2384	225					230					235					240	
2386	gat	gtg	gtt	tgg	cta	gaa	aaa	caa	gct	aaa	gaa	tgg	ctt	aaa	ttg		765
2387	Asp	Val	Val	Trp	Leu	Glu	Lys	Gln	Ala	Lys	Glu	Trp	Leu	Lys	Leu		
2388				245						250					255		

10/729,571

21

<210> 75

<211> 19

<212> DNA

<213> primer

<400> 75

tgatgcaaca aatggacga

19

<210> 76

<211> 18

<212> DNA

<213> primer

<400> 76

ttacaatttg agccattc

18

*invalid response - see ten 100m Error summary sheet*

**VERIFICATION SUMMARY**

PATENT APPLICATION: US/10/729,571

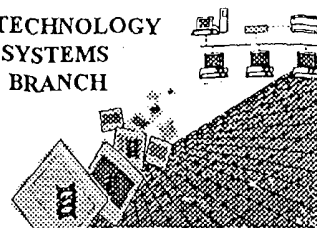
DATE: 12/22/2003

TIME: 10:11:24

Input Set : A:\SEQUENCE LISTING.txt

Output Set: N:\CRF4\12222003\J729571.raw

L:9 M:270 C: Current Application Number differs, Replaced Current Application Number  
L:10 M:271 C: Current Filing Date differs, Replaced Current Filing Date  
L:36 M:200 E: Mandatory Header Field missing, <220> Tag not found for SEQ ID#:1  
L:179 M:200 E: Mandatory Header Field missing, <220> Tag not found for SEQ ID#:3  
L:322 M:200 E: Mandatory Header Field missing, <220> Tag not found for SEQ ID#:5  
L:465 M:200 E: Mandatory Header Field missing, <220> Tag not found for SEQ ID#:7  
L:608 M:200 E: Mandatory Header Field missing, <220> Tag not found for SEQ ID#:9  
L:751 M:200 E: Mandatory Header Field missing, <220> Tag not found for SEQ ID#:11  
L:894 M:200 E: Mandatory Header Field missing, <220> Tag not found for SEQ ID#:13  
L:1037 M:200 E: Mandatory Header Field missing, <220> Tag not found for SEQ ID#:15  
L:1180 M:200 E: Mandatory Header Field missing, <220> Tag not found for SEQ ID#:17  
L:1323 M:200 E: Mandatory Header Field missing, <220> Tag not found for SEQ ID#:19  
L:1466 M:200 E: Mandatory Header Field missing, <220> Tag not found for SEQ ID#:21  
L:1610 M:200 E: Mandatory Header Field missing, <220> Tag not found for SEQ ID#:23  
L:1753 M:200 E: Mandatory Header Field missing, <220> Tag not found for SEQ ID#:25  
L:1896 M:200 E: Mandatory Header Field missing, <220> Tag not found for SEQ ID#:27  
L:2039 M:200 E: Mandatory Header Field missing, <220> Tag not found for SEQ ID#:29  
L:2182 M:200 E: Mandatory Header Field missing, <220> Tag not found for SEQ ID#:31  
L:2325 M:200 E: Mandatory Header Field missing, <220> Tag not found for SEQ ID#:33



## RAW SEQUENCE LISTING ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 10/729,571  
Source: IFWO  
Date Processed by STIC: 12/22/03

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.

PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

FOR CRF SUBMISSION AND PATENTIN SOFTWARE QUESTIONS, PLEASE CONTACT MARK SPENCER, TELEPHONE: 703-308-4212; FAX: 703-308-4221

Effective 12/13/03: TELEPHONE: 571-272-2510; FAX: 571-273-0221

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER VERSION 4.1 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:

<http://www.uspto.gov/web/offices/pac/checker/chkr41note.htm>

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.

Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom.

Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:

1. EFS-Bio (<<http://www.uspto.gov/efb/efs/downloads/documents.htm>> , EFS Submission User Manual - ePAVE)
2. U.S. Postal Service: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450
3. Hand Carry directly to (EFFECTIVE 12/01/03):  
U.S. Patent and Trademark Office, Box Sequence, Customer Window, Lobby, Room 1B03, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202
4. Federal Express, United Parcel Service, or other delivery service to: U.S. Patent and Trademark Office, Box Sequence, Room 4B03-Mailroom, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202

Revised 10/08/03



## Raw Sequence Listing Error Summary

### ERROR DETECTED

### SUGGESTED CORRECTION

SERIAL NUMBER: 10/729,571

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

- 1        Wrapped Nucleics  
    Wrapped Aminos     The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."
  
- 2        Invalid Line Length     The rules require that a line not exceed 72 characters in length. This includes white spaces.
  
- 3        Misaligned Amino  
    Numbering     The numbering under each 5<sup>th</sup> amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.
  
- 4        Non-ASCII     The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.
  
- 5        Variable Length     Sequence(s)        contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.
  
- 6        PatentIn 2.0  
    "bug"     A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s)       . Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.
  
- 7        Skipped Sequences  
    (OLD RULES)     Sequence(s)        missing. If intentional, please insert the following lines for each skipped sequence:  
                            (2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)  
                            (i) SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading)  
                            (xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)  
                            This sequence is intentionally skipped  
  
                            Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences.
  
- 8        Skipped Sequences  
    (NEW RULES)     Sequence(s)        missing. If intentional, please insert the following lines for each skipped sequence.  
                            <210> sequence id number  
                            <400> sequence id number  
                            000
  
- 9        Use of n's or Xaa's  
    (NEW RULES)     Use of n's and/or Xaa's have been detected in the Sequence Listing.  
                            Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present.  
                            In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.
  
- 10        Invalid <213>  
    Response     Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or is Artificial Sequence
  
- 11        Use of <220>     Sequence(s)        missing the <220> "Feature" and associated numeric identifiers and responses.  
                            Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section.  
                            (See "Federal Register," 00701/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)
  
- 12        PatentIn 2.0  
    "bug"     Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.
  
- 13        Misuse of n/Xaa     "n" can only represent a single nucleotide; "Xaa" can only represent a single amino acid